Recycling utterances: A speaker's guide to sentence processing

EWA DĄBROWSKA*

In recent years, there has been a growing consensus that speakers store large numbers of preconstructed phrases and low-level patterns, even when these can be derived from more abstract constructions, and that ordinary language use relies heavily on such relatively concrete, lexically specific units rather than abstract rules or schemas that apply "across the board". One of the advantages of such an approach is that it provides a straightforward explanation of how grammar can be learned from the input; and in fact, previous work (e.g. Dąbrowska and Lieven 2005) has demonstrated that the utterances children produce can be derived by superimposing and juxtaposing lexically specific units derived directly from utterances that they had previously experienced. This paper argues that such a "recycling" account can also explain adults' ability to produce complex fluent speech in real time, and explores the implications of such a view for theories of language representation and processing.

1. Introduction

An adequate cognitive theory of language must be able to account for its three basic properties: productivity, or our ability to produce and understand an indefinite number of novel utterances; learnability, or the fact that language learners are able to acquire this ability from a finite number of utterances that they are exposed to; and, last but not least, real-time processing: the fact that we are able to process complex messages very quickly, even in conditions which are far from ideal. In ordinary conversation, English-speaking adults produce, on average, 210 words per minute (Tauroza and Allison 1990); the average gap between turns is about 500 msec (Bull and Aylett 1998). Thus, within half of a second of the first speaker finishing her turn, the interlocutor has interpreted the speakers' utterance, decided on an appropriate response, formulated it, and launched an appropriate motor programme to begin his own contribution. Since conversational turns often overlap, it is clear that much of this processing happens while the second speaker is still listening. To make things worse, ordinary conversation often contains sentence fragments and false starts, as well as some ungrammatical sentences, and the acoustic signal is often distorted or partially inaudible[i]. Yet speakers and listeners are not fazed by this, and are able to communicate with each other reasonably effectively, with only occasional misunderstandings.

The generative tradition in linguistics has emphasized productivity, and failed on the other two criteria. Chomsky explicitly acknowledges that a generative grammar is unlearnable from the input – hence the need for an innate Universal Grammar (which, as Dąbrowska under review argues – is an extremely problematic concept). Most generative linguists have little interest in processing, which is regarded as part of "performance", and the linguistic models developed in this framework are computationally intractable and not easily accommodated in a processing theory (which would have to explain how people go from meaning to sound and vice versa).

Usage-based cognitive models, in contrast, can accommodate all three desiderata. They can explain productivity by appealing to constructional schemas (Langacker 1987, 2000) and offer a straightforward account of how children extract schemas from the language that they hear (see Dąbrowska 2000; Diessel 2004; Lieven et al. 2009; Tomasello 2003), thus fulfilling the learnability requirement. Last but not least, they can also explain how speakers process language in real time. This aspect of the theory has not been developed in much detail, however, and will therefore be the main focus of this paper.

In usage-based cognitive models (Barlow and Kemmer 2000; Bybee 2010; Goldberg 2006; Langacker 2000) linguistic knowledge is represented by a network of form-meaning pairings called constructions. Constructions vary in size and specificity: they can be simple or complex (i.e., consist of smaller subunits), and they can be concrete (phonologically specific) or schematic (partially or entirely underspecified). Idioms such as *kick the bucket* and formulaic phrases such as *I like it* are examples of complex concrete expressions; it is assumed that the latter are available as preconstructed units just like words of comparable frequencies.[ii] Complex schematic units include low-level schemas (or “frames with slots”) such as *I like NP* or *NP1 like NP2,* as well as more abstract units such as the transitive construction; crucially, even the highly abstract constructions are meaningful (cf. Langacker 1987; Goldberg 1995).

Figure 1 presents schematic diagrams of three types of complex units: a formulaic phrase, a low-level schema and an abstract schema. The diagram loosely follows the conventions developed by Langacker: the large boxes at the top and at the bottom represent semantic and phonological structure respectively, while the smaller boxes represent substructures within these complex units, and the vertical lines, symbolic relationships. The abbreviations ‘tr’ and ‘lm’ stand for ‘trajector’ (loosely, “subject”) and ‘landmark’ (“object”), the arrow represents a ‘process’, or temporal relation (see Langacker 1987 for a full exposition and explanation of these terms). Crucially, all three types of units have the same overall structure, and differ only in specificity (cf. Dąbrowska 2000; Langacker 1987); thus, the transition from formula to schema does not require translation into a different representational format, but merely loss of detail. To acquire a complex schematic unit, a learner must be able to do three things: (1) represent the phonological and semantic properties of utterances; (2) segment utterances and match chunks of phonology with chunks of semantic structure (e.g. work out that the phonological chunk /laik/ in Figure 1b corresponds to the relationship between the ‘liker’ and the ‘liked’); and (3) form slots by generalising over fillers attested in a particular frame (i.e., observe that the first slot in the frame *NP1 like NP2* is filled with a referring expression designating the “liker”, while the second slot corresponds to the “thing liked”). Note that relational words such as verbs or prepositions are effectively low-level schemas like the one represented in Figure 1b (see Dąbrowska 2009 for further discussion). We know that different languages lexicalise relational concepts very differently (Bowerman 1996; Dąbrowska 2004; Sinha et al. 1994), and hence children must be able to acquire such representations from the input available to them. One of the great advantages of the usage-based cognitive approach is that it can explain grammatical development without assuming any mental abilities over and above those which are necessary for lexical learning (see Tomasello 2003).

Thus, according to the usage-based approach, constructions are acquired by generalising over specific exemplars. Specific constructions are, from a developmental point of view, more basic than abstract ones. Later in this paper, I will argue that this is also the case from the point of view of adult production and comprehension: speakers prefer more concrete units, and use them whenever they can. Finally, the usage-based approach assumes that constructions become entrenched through repeated use: every use of a particular unit, whether in comprehension or production, strengthens it and increases the likelihood of its being used in the future. Thus, other things being equal, more frequent expressions are more likely to be available as units, and more likely to be accessed during language processing.

Figure 1: Types of complex units
(a) a formulaic (fixed) phrase
(b) a low level schema (“frame with slots”)
(c) a general schema (“abstract construction”)

The paper is organized as follows. In the next section, I discuss some language acquisition studies which demonstrated that the vast majority of young children's multiword utterances can be easily traced back to precedents in the language they had experienced, and that their syntactic productivity is best explained by assuming that they recombine lexically specific chunks derived from their experience. In Section 3, I review evidence that lexically specific chunks are also used by adults. The following two sections describe a case study examining the extent to which a particular adult's conversational speech relies on such units. Section 6 provides a brief outline of a usage-based model of sentence production, and the concluding section discusses the implications of such view of production for theories of language structure.

2. How children build sentences (and learn grammar in the process)

Dąbrowska and Lieven (2005) and Lieven et al. (2009) describe an application of the usage based approach to child language. Both studies begin with the observation that early child language is highly formulaic (cf. Lieven et al. 1997, 2003; Tomasello 2000), and attempt to explain children's grammatical productivity in terms of lexically specific units derived from the input and simple operations on these units. Both studies analysed data for four children who had been recorded for 30 hours each at ages 2 and 3. Each of the resulting eight corpora was divided into two parts: the test corpus, which comprised the last, or the last two, transcripts, and the main corpus, which comprised the remaining transcripts. The child's multiword utterances in the test corpus were then "traced back" to lexically specific units attested in the main corpus. It was reasoned that if this was possible, then the inventory of lexically-specific units derived from the main corpus could be said to capture the children's syntactic knowledge at this stage of development.

The traceback analysis assumed two kinds of lexically specific units, fixed phrases and frames with slots. A **fixed phrase** was defined as a recurrent form-meaning combination; it was assumed that it was available to the child as a unit if it was attested at least two times in the main corpus, in either the child's or an adult's speech. (A frequency criterion of 2 may seem rather unrealistic: I return to this issue later. For now, the reader should bear in mind that the main corpora are relatively small samples of the children's linguistic experience: about 1% in the case of the two-year-olds and about 0.5% in the case of the three-year-olds.) For instance, the following utterances from the Thomas 3 main corpus were regarded as evidence that the expression *some milk* was available to the child as a preconstructed unit. (Note that these are examples: there in total 13 tokens of *some milk* in the main corpus.)

(1) *CHI: *cornflake-s with <u>some milk</u>.*

*CHI: *I want <u>some milk</u>.*

*CHI: *I got <u>some milk</u> and orange juice here.*

*MOT: *<u>some milk</u> .*

*MOT: *well you can have <u>some milk</u> at bedtime .*

*MOT: *would you like <u>some milk</u> on them , Thomas?*

Frames are recurrent form-meaning combinations containing a "slot" into which novel material can be inserted. To qualify as a frame, an expression had to be attested in the main corpus at least twice, with at least two different fillers in the slot. For instance, the sentences in (2) are regarded as evidence for a *buy some N* frame, where the N slot is filled by any expression designating a quantity of things that can be bought. The main corpus contained 28 tokens of this phrase. The sentences in (3) are instances of the *do you want to VP* frame (attested 47 times), where the VP slot can is filled by an expression designating the proposed activity. For simplicity, I use traditional syntactic labels such as N and VP to identify the slots, but it is assumed that the categories are defined in semantic terms and are frame-specific (at least initially): that is to say, the class of fillers that are inserted into the slots in these two frames are "things that can be bought" and "activity you could do".

(2) *CHI: *I buy some jelly+baby-s.*

*CHI: *buy some sweet-s.*

MOT: well we must buy some more cheese from the supermarket.*

MOT: well you usually buy some little jelly # worm-s.*

MOT: we set off in the morning for a quick walk to buy some milk and we went via Lynne and Ben and Charlotte-'s house ,, did-'nt we?*

*MOT: *you-'re go-ing to buy some orange+juice and milk and post some letter-s?*

(3) *CHI: *do you want to kiss this?*

*CHI: *do you want to read my post+box?*

*MOT: *do you want to do it ?*

*MOT: *do you want to press the button?*

*INV: *ah do you want to show me?*

*MOT: *do you want to get to your toybox?*

Lexically specific units can be combined using two "usage-based" operations, juxtaposition and superimposition. **Juxtaposition** involves the simple concatenation of two independent units; the relationship between the units is paratactic rather than syntactic, and they can occur in either order. This occurs most often with vocatives, parentheticals, and some adverbials. For example, if we juxtapose the fixed phrase *where are you* and the noun *baby*, we get *where are you, baby?* or *baby, where are you?*; juxtaposing *where's it gone* and *now* results in *where's it gone now?* or *now where's it gone?* **Superimposition** involves the combination of a frame with another chunk in such a way that the corresponding elements are "fused"; this happens at the phonological level and the semantic level at the same time. The simplest case is illustrated in Figure 2, where the noun *pteranodon* elaborates the noun slot in the frame *where's the …*.. (The resulting utterance, *where's the pteranodon?* was produced by one of the three-year-olds.) Figure 3 shows the superimposition of two frames, *get them AP* and *get NP ready*; the corresponding elements are connected with dotted lines. (For the sake of simplicity, I use a simplified notation which does not show semantic and phonological structure separately, but it should be remembered that all the units shown are form-meaning pairings, and that the operations apply simultaneously at both levels.) The two frames are "superimposed" in such a way that *them* in the first frame elaborates the NP slot in the second frame, and the element *ready* from the second frame elaborates the AP slot in the first frame; the result of the superimposition is the novel expression *get them ready.* Superimposition can only apply when the filler has the properties required by the slot: an NP slot can be filled only by an expression referring to a "thing" (in the technical cognitive grammar sense – see Langacker 1987), an AP slot can only be filled by a word describing a property, and so on.

Note that the expression *get them ready* could also be derived by superimposing *them* over the NP slot in *get NP ready* or by superimposing *ready* over the AP slot in *get them AP.* However, Dąbrowska and Lieven assumed that the child would always use the largest available chunks – in this case, *get them AP* and *get NP ready.* This assumption was made for two reasons. First, it was hypothesized that using overlapping chunks would facilitate superimposition, since it would make it easier to see how the chunks "fit together". Secondly, using larger chunks eliminates many errors of commission, particularly in cases involving government and agreement. For instance, if the child used the *get NP ready* frame and combined it with the pronoun *they*, the result would be ungrammatical, while using a larger chunk such as *get them AP* results in a grammatical structure.

Figure 2: The superimposition of *where's the N* and *pteranodon*

Figure 3: Superimposition of *get them AP* and *get AP ready*

To get a sense of how utterances can be constructed using these units and operations, consider the example in Figure 4, which shows how one of the target utterances, *Squirrel, do you want to buy some milk* (produced by Thomas at age 3) can be assembled by combining four chunks: *do you want to VP* (attested 47 times), *buy some N* (attested 28 times), *some milk* (attested 13 times), and *squirrel* (attested twice). The derivation involves two superimpositions and one juxtaposition; notice that these operations can apply in any order.

Figure 4: Derivation of *Squirrel, do you want to buy some milk?*

Dąbrowska and Lieven (2005) and Lieven et al. (2009) found that, at age 2, the majority of the children's utterances were either zero-operation utterances (exact repetitions of something either the child or the mother had said at least twice before in the main corpus) or single-operation utterances (minimal modifications of previously experienced utterances). This was also true of three of the four children at age 3;0, although they were using more 2- and 3-operation utterances and in one child's case also some 4-operation utterances. At both ages 84-100% of the children's questions can be derived by using lexically-specific units attested in the main corpora. The remaining utterances were "fails", i.e., unsuccessful derivations. Although fails could, in principle, be regarded as evidence for more abstract knowledge, there is good evidence that this is not the case. Dąbrowska and Lieven (2005) point out that the "fails" are no more complex than the successful derivations, yet much more likely to be ungrammatical, and argue that they are probably speech errors, one-off innovations, or simply sampling artefacts (the main corpus comprised only 0.5-1% of the children's linguistic experience, and hence many expressions which were unattested in it could in fact have been available to the children as units). We can conclude, then, that young children's syntactic creativity can be explained by assuming that they "recycle" memorized chunks.

3. Holistic processing in adults

We have seen that young children's ability to produce sentences can be explained by assuming that they retrieve lexically specific chunks from memory and combine them using the operations described earlier. Could such a mechanism also account for adult productive abilities?

Following early demonstrations of the importance of preconstructed phrases by Sinclair (1991) and Pawley and Snyder (1983), several researchers have tried to determine what proportion of texts is made up of such units. One of the best-known studies by Erman and Warren (2000) found that prefabs constitute 59% of spoken discourse. However, this figure is probably an underestimate, since Erman and Warren counted only conventionalized word combinations, i.e., combinations that contained some sort of idiosyncrasy. Altenberg (1998) counted all recurrent word in the London-Lund corpus, and found that over 80% of the word tokens in the corpus were part of such combinations. Note that even this figure could be an underestimate, since their corpus was quite small (less than half a million words), and Altenberg only counted recurrent contiguous combinations (in other words, he did include prefabs with slots).

On the other hand, as many critics pointed out, the fact that a combination of words recurs does not necessarily entail that speakers store it. Recurrent word combinations consisting of frequent words such as *of the* or *and it* are not particularly interesting linguistically, and could arise purely by chance. However, there is growing evidence that speakers store large numbers of preconstructed phrases, including completely regular phrases (Cowie 1998; Ellis 1996; Siyanova-Chanturia and Martinez 2014; Weinert 2010; Wray 2002), and that these facilitate sentence comprehension and production: for instance, they are read faster (Arnon and Snider 2010; Conklin and Schmitt 2008; Tremblay et al. 2011; Siyanova-Chanturia et al. 2011), produced more quickly (Tremblay and Tucker 2011) and more fluently (Dąbrowska 2008a; Kuiper 1996) and remembered better (Dąbrowska et al. 2009; Tremblay et al. 2011) than non-formulaic chunks.

Most researchers, however, continue to regard formulaic language as either a developmental phenomenon – a stage that learners go through before they become fully productive with the rules of a language – or as special mode of processing which complements the normal (i.e., analytic) mode. Thus, while juxtaposing and superimposing memorized chunks may work for young children, who produce short and relatively simple sentences, it does not appear to be a viable method of sentence production for adults, except in some special circumstances. Since the number of possible word combinations increases exponentially with sentence length, adults couldn't possibly store enough units, or retrieve them when needed during real-time language processing.

Or could they? Human long-term memory has a vast storage capacity. Some researchers have even argued that, for all practical purposes, its storage capacity can be assumed to be infinite (cf. Barsalou 1992: 148; Schneider and Bjorklund 1998: 469; Standing

1973), and that encoding into memory is "an obligatory, unavoidable consequence of attention" (Logan 1988: 493). This is perhaps most striking with visual stimuli. In a classical study by Shephard (1967), participants were shown each of 612 pictures for six seconds. When tested later using a two-alternative forced-choice recognition task, they were able to choose the correct picture 98% of the time. Standing (1973) used a similar method with much larger stimuli sets; but even when presented with 10,000 pictures, participants correctly recognized the seen pictures 83% of the time.

Memory for verbal stimuli appears to be less accurate than for pictures, but it is still impressive. A series of rigorous experiments recently conducted by Gurevich et al. (2010) demonstrates incidental verbatim memory for sentences in adults after a single exposure. Participants listened to one of two versions of a story and were then given a surprise recognition test. The stories were constructed in such a way that each sentence in version 1 corresponded to a nearly synonymous sentence in version 2: for instance, *It's bedtime," said Jesse* corresponded to *"Time for bed," said Jesse.* The sentences from the other version of the story were used as foils on the recognition task. Gurevich et al. found that, after a single exposure to the story, participants recognised sentences at above chance levels (60-73% correct), and were able to reproduce some of the sentences verbatim, even when tested a week later.

Another relevant source of evidence comes from studies which examined the corpus frequency of conventionalized multiword units such as idioms and proverbs. Because such units are, by definition, (at least partly) idiosyncratic, they must be stored, yet their text frequencies are often very low. For instance, Moon (1998) examined the occurrence of conventional multi-word units extracted from a dictionary of idioms in an 18 million-word corpus. She found that 70% of the units in her data set had a frequency of less 1 per million, and 40% did not occur at all in the corpus. The later included expressions such as *hang fire,* *kick the bucket, speak for yourself,* and *one man's meat is another man's poison* – which one could reasonably expect most native speakers to know.

Another striking feature of human memory is that it is content addressable, that is to say, we can retrieve a whole stored pattern when a part of the pattern is available as cue: for example, a semantic representation of a phrase can be used to retrieve the whole symbolic unit, or a part of a phrase can be used to retrieve the whole phrase. As a result, retrieval is very efficient (we do not have to serially search through all our memories to retrieve a specific one, as most computers do), and retrieval speed is largely independent of the amount of the stored data (in fact, for domains that are richly interconnected, the more a person knows about the domain the better they are at encoding and retrieving information about it).

4. The data

So the idea that sentence production in adults, as in children, depends largely on retrieval of preconstructed chunks is perhaps not as outlandish as one might at first think. However, a researcher wishing to conduct a "traceback" analysis of a sample of adult utterances faces some practical problems. Adult speakers have experienced vast amounts of language. In ordinary conversation, speakers produce about 210 words per minute (Tauroza and Allison 1990). Assuming 6 hours of exposure to language per day, adult language users experience over half a million words (210×60×6×7=529,200) every *week* of their lives. Thus, to conduct a traceback analysis of adult speech, we need a vast corpus, ideally containing the speech of a single individual – and most large corpora consist mainly of written language, typically collected from many different users. In this respect, developmental corpora of child speech are rather unique in that they contain data from a small number of interlocutors recorded in informal, family settings.

Note, however, that developmental corpora also contain adult speech – the speech addressed to the child who is the object of study. The largest developmental corpus currently available is the Thomas corpus collected by Lieven et al. (2009) and available from the CHILDES database (MacWhinney 2014). Thomas was recorded for five hours a week from age 2;0 to 3;0, and for one hour a week from 3;0 to 5;0. The corpus contains transcripts of 379 hours of informal conversation. This includes 1,816,691 words produced by the mother (as well as 528,332 words produced by the child, and 164,689 words produced by other speakers). It is thus larger than any other sample of a single adult's conversational experience available. This corpus was therefore used in the analysis described here. The test corpus comprises a random sample of 100 multiword utterances produced by the mother (henceforth M) in the last recording, which was made when Thomas was 4;11.20. The main corpus comprises all the utterances in all preceding transcripts.

Before we proceed, it is important to examine some of the properties of this corpus. It is well known that child-directed speech (CDS) differs from adult-directed speech in a number of respects: it is lexically less diverse, contains shorter, grammatically simpler utterances, is more repetitive, and tends deal with a more restricted range of topics (Cameron-Faulkner et al. 2003; Newport et al. 1977; Snow 1986; Stoll et al. 2009). This raises an important question: can Thomas' mother's speech be regarded as a representative example of adult conversation?

To address this point, we should first note that all the studies just mentioned investigated language addressed to children aged between one and three, whereas the test corpus which will be analyzed here comes from a recording made when Thomas was almost five. Furthermore, it if often asserted that grammatical development is complete by age four or five[iii]; and if four-year-olds know all there is know about (basic) grammar, then presumably adults' speech to them does not differ grammatically from ordinary adult directed speech. Actually, however, children's grammars continue to evolve until well into adolescence or even later, mainly a result of experience with written language (Berman 2004, 2007; Hoffer 1994; Nippold 1998; Nippold et al. 2005). Thus, while language addressed to a five-year-old is likely to be more complex than that addressed to younger children, it could still be grammatically and/or lexically simpler than adult-adult conversation. To ensure that the sample analysed here is representative of ordinary adult speech, I conducted two analyses, one comparing the gammatical complexity and lexical diversity of the sample of M's speech with normative data for adults, and a second which directly compares it with samples of adult conversation.

Nippold et al. (2005) have identified two measures which are particularly useful in measuring grammatical complexity in later stages of grammatical development: average length of speakers' T-units (i.e., the main clause plus any subordinate clauses attached to it) and clausal density, or subordination index (the average number of clauses per T-unit). As shown in table 1, both of these measures increase slowly but steadily from age 8 until at least age 20. The mean length of T-unit in the mother's sample is 8.60 and its clausal density of 1.30. This is slightly below the average figure for her age group (20+), but well within the range.

To assess the lexical diversity of the mother's sample, I compared it with the findings reported in Johansson's (2008) study. The traditional measure of lexical diversity, the type-token ratio (TTR), is known to be highly dependent on sample size, with larger samples giving considerably lower TTR values. For this reason, Johansson used a new measure called D, which has been shown to be more accurate (MacWhinney 2014; Richards and Malvern 1997). The mean value of D for adults in the Johansson study was 61.11 for narratives and 81.12 for expository speech. (Unfortunately, Johansson does not provide information about the range or the standard deviation). The corresponding figure for the mother (computed using the VOCD programme in CLAN) was 101.60 – well above the adult mean.

Table 1: Mean length of T-unit (in words) and clausal density in conversation (Nippold et al. 2005)

Age	Length of T-unit			Clausal density		
	Mean	SD	Range	Mean	SD	Range
8	6.74	0.86	4.42–8.44	1.18	0.09	1-1.37
11	7.31	1.62	3.67–10.56	1.25	0.22	1-1.75
13	6.88	0.93	5.56–8.52	1.17	0.09	1-1.42
17	8.33	1.27	5.83–10.32	1.30	0.11	1.08-1.56
25	9.86	2.08	6.00–13.44	1.39	0.21	1.08-1.82
44	9.56	2.01	6.88–15.16	1.38	0.19	1.12-1.77

As a final test, I directly compared the mother's sample with transcripts of adult conversations taken from the Santa Barbara Corpus of Spoken American English (Du Bois et al. 2000-2005). Data from three transcripts, representing seven speakers, were used in the analysis. The transcripts were chosen to represent a range of different conversational situations: three friends, Marilyn, Roy and Pete, chatting while cooking a meal together (SBC005), a married couple (Darryl and Pamela) discussing the possibility of life after death (SBC003) and, as an example of task-based talk, two teenagers (Kathy and Nathan) going over a math test (SBC009). Samples of approximately 1100 words were extracted for each speaker. Table 2 below presents the results of the analysis using the two grammatical measures introduced above (mean length of T-unit and clausal density) and two measures of lexical diversity (TTR and D). As shown in the table, M's speech is indeed comparable to that of the other seven speakers.

Table 2: Comparison of the mother's sample with adult-adult conversations

Speaker	Clausal density	Mean length of T-unit	TTR	D
Kathy	1.20	6.78	0.22	59.88
Nathan	1.34	5.43	0.27	69.35
Darryl	1.37	7.71	0.39	90.71
Pamela	1.41	7.68	0.36	91.67
Marilyn	1.14	5.93	0.43	108.82
Roy	1.30	7.70	0.41	113.14
Pete	1.13	6.05	0.45	91.39
Mean	**1.27**	**6.75**	**0.36**	**89.23**
M	1.30	7.40	0.40	101.08

I conclude, therefore, that the corpus can be regarded as a representative sample of adult conversational speech. Ideally, of course, one would use a suitably large sample of an adult speaker's conversations with other adults, but until such a corpus is available, we will have to make do with what we have. It should also be pointed out that the bulk of the main corpus (56%) contains transcripts of recordings made when Thomas was less than three years old, and is thus likely to contain less complex speech. This makes it more difficult to find lexically specific precedents for the mother’s utterances; thus, the structure of the corpus is actually biased against the hypothesis that adult speech can be explained in lexically-specific terms.

5. Case study

This section describes a case study conducted in order to assess the role of preconstructed chunks in spontaneous adult production. The test corpus (i.e., the target utterances) comprises a random sample of 100 multiword utterances produced by M in the last recording, which was made when Thomas was 4;11.20. The target utterances were analysed using the same procedure as in the child study, with one exception: frames could contain up to two slots rather than just one, as in the child study. Examples of two-slot frames include the following: *NP could do without NP* (instantiated by *I could do without them, Thomas*), *remember TIME when S* (e.g. *remember before when I asked you if you'd enjoyed your lollipop from Marks+and+Spencers*), and *we have to VP and VP* (e.g. *we have to change direction and go over the bridge*).

5.1 General summary of the results

Figure 5 provides information about the number of operations needed to construct the target utterances. As we can see, about a third of the utterances produced by M were zero or one operation utterances, i.e., either exact repetitions of a phrase attested at least twice in the main corpus, or utterances involving superimposition or juxtaposition of two previously attested units. Another third of the target utterances can be assembled by using two to four operations. Of the remaining utterances, 16 require 5-13 operations; 8 are lexical fails (that is to say, they contain a word which was not attested at least two times in the main corpus); and 10 are syntactic fails.

Figure 5: Number of operations needed to derive M’s utterances

The component units vary in length from 1 to 7 words. (For the purpose of this analysis, slots were counted as one word; contracted forms such as *she's* and *didn't* were also counted as single words). The frequency distribution of units of different length is shown in Figure 6. As we can see, the majority of multi-word units were relatively short (2-4 words). The frequency distribution is similar to that observed in Moon's (1998) study of fixed expressions and idioms in English.[iv]

Figure 6: Frequency distribution of component units of different length

Table 3 shows the frequency distributions of all units used in the traceback. (Note that the figures are for tokens, not types: thus, the frame *and S* was counted 7 times, since it occurred in seven different sentences.) The columns in the table correspond to different types of units. Columns 2 and 3 provide the figures for lexically specific units, i.e. single words without slots (*book, ouch*) and formulaic phrases (*a nice cup of coffee*). Columns 4 and 5 provide the same information for formulaic frames; these are subdivided into words with slots (frames consisting of a single word and one or more slots, e.g. a*nd S, terribly ADJ*; note that traditionally such units would be considered single words) and phrases with slots (e.g. *you won't VP, will you? I think it might be NP*). Each row in the table represents a single frequency band. For instance, the first row tells us that 20 of the 311 units used in derivations, i.e. 6% of the total, occurred only twice in the main corpus; of these, 1 was a single word, 4 were fixed phrases, 1 was a word+slot(s) combination, and 14 were phrases with slots.

Table 3: Frequency distribution of all units used in the traceback analysis

Frequency	Single words	Fixed phrases	Words w/ slot(s)	Phrases w/ slot(s)	Total	% total units
2	1	4	1	14	20	6
3	0	5	1	9	15	5
4	2	7	1	13	23	7
5	4	5	1	3	13	4
6	1	4	0	5	10	3
7	1	3	1	5	10	3
8	0	2	0	4	6	2
9	1	2	2	4	9	3
10	1	0	2	2	5	2
11-20	3	9	3	14	29	9
21-30	5	6	0	8	19	6
31-40	1	6	0	6	13	4
41-50	4	1	0	5	10	3
50+	24	31	29	45	129	41
Total	48	85	41	137	311	100
% total	15	27	13	44	100	

There are several observations that one might make about these figures. First, most of the units used in the derivation are phrasal: fixed phrases and phrases with slots together account for 71% of all units. Secondly, the lower frequency bands (2-10) contain more phrasal units than single words with or without slots. This is not surprising, since individual words typically occur in a number of different phrases. Thirdly, the majority of the units are are relatively frequent: only 23% have a frequency between 2 and 5, and a relatively large proportion (41%) were very frequent, i.e., attested 50 times or more in the main corpus. This will be important in Section 5.4, where we discuss the implications of using different frequency thresholds for identifying ready-made units. Finally, we may note that invariant units (fixed phrases and single words) account for 42% of the units and one-slot frames for 48%. Only 8% of the component units are two-slot frames.

One of the purposes of the analysis conducted here was to investigate the processing advantages of using prefabricated chunks. Researchers studying formulaic language generally assume that retrieving an item from memory is computationally less demanding than assembling it from minimal components – but just how much processing effort is saved? We can obtain a rough estimate by comparing the number of operations required to assemble the target utterances using larger chunks identified in the main corpus with the number of operations necessary to assemble the same utterances "from scratch", that is, using chunks consisting of single lexical items. The mean length of the utterances analysed here was 7.4 words; the average number of operations per utterance was 2.3. A seven-word utterance assembled "from scratch" requires six operations and an eight-word utterance, seven operations. Thus, using chunks could save the speaker/listener about two-thirds of the processing effort.

It must be stressed that the above estimate is a very rough calculation based on simplifying assumptions which could be incorrect. It assumes that all chunks are equally easy to process, which may or may not be the case. Dąbrowska and Lieven (2005) argued that language learners prefer larger chunks, since the overlapping parts make it easier to see how they should be put together: unifying such chunks is thus similar to completing a jigsaw puzzle in which each piece has a picture of part of the adjacent piece attached to it. It could be argued, however, that larger chunks are more difficult to retrieve and manipulate. It is also likely that chunk frequency plays a role, in that more frequent and hence more entrenched chunks presumably incur a lower processing cost; the above estimate, however, does not make allowances for this (On the other hand, as discussed in Section 5 below, smaller chunks may act as retrieval cues for larger ones, which may partially offset the difficulty of manipulating larger chunks.) Secondly, the estimate given above assumes that juxtaposition and superimposition involve the same amount of effort. However, it could be argued that superimposition is more effortful because it requires attention to the internal structure of chunks and lining up corresponding elements, while juxtaposition does not require any such processes. Thus, the figure given in the preceding paragraph must be treated with some caution; it does, however, give us a sense of how much processing effort we could *potentially* save by using larger chunks.

5.2 Some examples

To give the reader a sense of how the adult traceback works in practice, I show how four fairly complex utterances can be derived by juxtaposing and superimposing chunks. I then provide a full list of all the "fails" (i.e., utterances which could not be derived from lexically specific units attested in the main corpus) and discuss them briefly.

5.2.1. Example 1: The cover for the box is still there it's still in the box look.

Example 1 consists of three clauses: (1) *the cover for the box is still there,* (2) *it's still in the box* and (3) *look* juxtaposed to form a single utterance.[v] The first clause can be assembled by superimposing *the cover for the N* and *the box,* which yields *the cover for the box;* this, in turn, can be superimposed over the NP slot in the frame *NP is still there.* The other two clauses are assumed to be available as units. Thus, the target utterance by combining the five chunks:

NP is still there (3)

the cover for the N (2)

the box (50+)

it's still in the box (2)

look (50+)

and requires four operations. (The numbers in parentheses indicate corpus frequency; 50+ means that the chunk occurred more than 50 times.)

This analysis assumes that any chunk with a corpus frequency of two or more is available as a ready-made unit. Is this a realistic assumption? It is important to bear in mind that the transcripts available in the corpus constitute about 0.26% of M's linguistic experience.[vi] Thus, the *average* frequency of lexical chunks with a corpus frequency X in M's experience is roughly 385 times X. In other words, if we had to estimate how often she experienced the chunks *the cover for the N* and *it's still in the box*, our best guess would be 770 times. Given that the Gurevich et al. study described earlier demonstrated some verbatim memory for sentences after a *single* exposure, it does not seem unreasonable to assume that a chunk which had been experienced 770 times is available as a preconstructed unit.

Of course, we have no way of knowing for sure how many times M had encountered a particular lexically specific unit: the two recorded tokens may have been the only two she ever experienced, and we were just lucky to catch them in the recording. It is important to note, therefore, that nothing crucial hinges on the use of the specific chunks enumerated earlier. If we raise the frequency threshold to 3, *the cover for the N* and *it's still in the box* are no longer available as preassembled units, but the target utterance can be assembled using different chunks: *cover for NP* (which is attested 7 times in the corpus) and *it's still in the N* (attested 4 times). If the frequency threshold is raised to 4, *NP is still there* is no longer available and we have to use the two-slot frame *NP is still LOC,* which is attested 8 times; and so on. Thus, as the frequency threshold is raised, more abstract units must be used, which results in more complex derivations; however, the target can still be assembled.

5.2.2. Example 2: I don't think we really need a train track, do you?

Note that in this utterance, the pronoun in the main clause doesn't match the pronoun in the tag. Thus, the utterance could not be derived using traditional tag rules. However, the frame *I don't think S do you?* is well attested in the corpus (117 tokens), and the target utterance can be easily assembled by combing this frame with two other lexically specific units: *I don't think we really need NP* (attested twice) and *a train track* (attested 29 times). If we raise the frequency threshold, *I don't think we really need NP* is no longer available and the target utterance will have to be constructed using more frequent chunks: either *we really need NP* (attested 3 times) or *NP really need NP* (attested 7 times).

5.2.3. Example 3: You're not going to tell me what happened when you went to Luke's house, Thomas.

This quite complex utterance can be assembled in 4 steps using the following chunks:

you're not going to VP (91)
tell me what happened when S (3)
when you went to POSS house (4)
Luke's N (7)
Thomas (50+)

Again, there are various alternative derivations. Let us consider just one part of this utterance: the subordinate clause *when you went to Luke's house.* The simplest derivation involves superimposing two complex frames, *when you went to POSS house* + *Luke's N*. However, there are a number of other possibilities. These include the following (among others):

when S (50+) + *you went to POSS house* (7) + *Luke's N* (7)

when S (50+) + *you VP* (50+) + *NP went to POSS house* (45) + *Luke's N* (7)

when S (50+) + *you went to NP* (50+) + *POSS house* (50+)+ *Luke's N* (7)

when you went DIR (41) + *to NP's house* (50+) + *Luke* (50+)

when NP went to NP's house (16) + *you VP* (50+) + *Luke* (50+)

5.2.4. Example 4: It was quite annoying really because you thought oh great I've found ten pee but there was no way you could lift it up.

All of the examples considered so far were relatively simple derivations involving 2-4 operations. In this final example, we consider an utterance which was somewhat problematic, requiring 12 operations. (Another target utterance contained a long list of nouns and required 13 operations. However, it was not syntactically interesting and will not be discussed further.) The utterance can be assembled by juxtaposing and superimposing the following chunks:

It was AP because S (8)

(NP BE) quite AP really (11)

annoying (15)

because you thought S (7)

(NP thought) oh EXCL S (4)

great (50+)

I've found NP (49)

ten pee (3)

but S (50+)

there BE no way S (4)

there was NP (50+)

you could VP (50+)

lift it up (50+)

Although the derivation is quite complex, it is worth observing that using preconstructed chunks still saves half of the processing effort: if the utterance had been assembled from word-sized units, the derivation would involve 24 steps.

5.3 Unsuccessful derivations ("fails")

Eighteen of the target sentences could not be derived from lexically specific units; a complete list of these is given in Table 4, with the problematic parts printed in bold. The first eight of these "fails" are purely lexical: that is to say, one of the words used in the utterance was not attested in the relevant sense, or at all, in the corpus, or it was attested only once and hence did not meet the frequency criterion. As can be seen from the table, all of these are perfectly conventional uses of ordinary English words such as *stroke, explosion,* or *ease*, and we can be confident that M has both heard and used these words many times before; the fact that they are not attested in the main corpus is simply an accident of sampling (recall that the sample only contains approximately 0.26% of her linguistic experience). But this figure puts the syntactic fails in perspective: the fact that there is no lexically specific match for a particular part of the target utterance may mean that it is produced using more general rules –

but it could also have been produced using lexically specific units which just happen not to be represented in the main corpus.

The remaining ten utterances are classified as syntactic fails. In three of them (9, 10 and 11), verbs which were only attested in a transitive frame in the main corpus are used intransitively, or vice versa; hence, although the problem is syntactic in the sense that it involves combinatorial properties, it is also lexical in the sense that the properties are associated with specific verbs. Sentence 12 is problematic because of the expression *on her mobile phone.* The nearest matches for this are *on the phone* and *on the telephone,* with no variation in the determiner slot. The sentence could be derived if we relaxed the variation criterion for frames with slots, which requires the frame be attested with at least two different fillers. The next two sentences, 13 and 14, involve slightly idiomatic uses which could be accommodated if we allowed for three-slot frames. The remaining four involve more complex sentences with a non-finite subordinate clause (15), a cleft (16), and relative clauses (17 and 18). Some of these could be accommodated if we assumed more abstract frames (e.g. *what NP MODAL do TIME is VP* for 16), but this seems to be stretching the definition of a lexically specific unit too far. The relative clauses are particularly problematic since they require substitution of a filler containing a gap – although in principle such constructions can be accommodated in terms of lexically specific templates (see Dąbrowska 2008a).

To summarize: the "traceback" method developed by Dąbrowska and Lieven (2005) worked very well for over 80% of M's utterances. Of the eighteen "fails", eleven involve specific properties of individual words, and three could be accommodated if the criteria for formulaic frames were slightly relaxed. The remaining four are more problematic and may constitute evidence for more abstract knowledge.

Table 4: Unsuccessful derivations

Utterance	Reason for fail
1. Thomas don't **fib** come on I haven't time.	only 1 token of *fib* as a verb in the main corpus
2. but not for quite a few because he's had a **stroke** now.	only 1 token of *stroke* (in the relevant sense) in the main corpus
3. can you not have a quiet **explosion**?	only 1 token of *explosion* in the main corpus
4. do you know what that nice music **support** [?] is?	only 1 token of *support* in the main corpus
5. I said that's a new one **on** me.	no occurrences of *on* in the relevant sense in the main corpus
6. something explodes and you can't sort of **plan** when it's going to happen.	no tokens of *plan* as verb in the main corpus
7. that's it gently **ease** it out.	only one token of *ease* as verb in the main corpus
8. tracks and **explosions** and trees falling down.	only 1 token of *explosion* in the main corpus
9. Thomas please if you do that I'm **stopping playing**.	*stopping* only used in intransitive frames in main corpus
10. well I wasn't **lifting** I was helping you into your chair I thought I felt something.	*lift* used only in transitive frames in main corpus
11. well don't **tease** today Thomas ,, will you …	*tease* used only in transitive frames in main corpus
12. I think she's obviously talking to someone **on her mobile+phone** ,, isn't she?	the only matches are *on the telephone, on the phone;* no variation in the determiner slot
13. and somebody else **is with us**.	requires a 3-slot frame (*PERSON BE with PERSON*) or matching *with us* with LOCATION in *NP is LOCATION*
14. what's the problem **why is there a caution sign up**?	requires a 3-slot frame (why *BE there* NP LOC)
15. … and the police+helicopter just **stayed above shining a light** for them.	no match for gerundive clause

16. **what we'll do after lunch** is brush your teeth.	requires a frame with four slots (*what NP MODAL do TIME is VP*)
17. and **the number of people I used to see** going to pick up that ten pee.	no match for relative clause
18. it's just **a model that somebody didn't want**.	no match for relative clause

5.4 Changing the frequency threshold

The analysis presented above assumed that units attested at least twice in the main corpus were available to M as ready-made chunks. This frequency threshold is admittedly arbitrary, and the reader may well wonder whether it is psychologically realistic. As argued earlier (see Section 5.2.1) the true frequency of most of these chunks in M's linguistic experience is likely to be considerably higher, but for any specific chunk, it is possible that the two occurrences attested in the corpus are the only tokens she had experienced. It is important, therefore, to determine to what extent the results of the traceback depend on the frequency assumption: in other words, how they would change if we assumed a higher frequency threshold, for instance 3, 4 or 5.

First, let us consider what happens when we raise the frequency threshold to 3. As shown in Table 3, there are 20 units with a frequency of 2 in the main corpus. These were used in the derivation of 16 utterances (4 derivations involved two units with a frequency of 2). Thus, 16 target utterances will be affected if the frequency is raised to 3. Of these, 5 can be derived using the same number of operations but smaller (and hence more frequent) chunks; 9 can be derived but require an additional operation; and 2 are lexical fails. Both of the latter involve proper names (*Lucy* and *Oxford*). There are no syntactic fails.

If we raise the minimum frequency threshold from 3 to 4, 13 sentences are affected. Of these, 1 requires the same number of operations with smaller chunks, 8 require one additional operation, and 4 requires two additional operations. There are no additional fails.

Finally, if we raise the minimum frequency threshold from 4 to 5, 19 sentences are affected. Of these, 4 require the same number of operations, 12 involve one additional operation, 1 is a lexical fail involving the relatively infrequent word *clatter,* and 2 are syntactic fails. The syntactic fails include a sentence containing a passive with an oblique prepositional phrase (*there it's* ***attached to the other end of the train+track***) and a nominalized verb followed by a 'dangling' preposition (*what did you* ***have a play with*** *at Luke's house?*) – both somewhat unusual structures.

Table 5: Traceback results under different frequency thresholds

Derviation result	Frequency threshold			
	2	3	4	5
Successful	82	80	80	77
Lexical fail	8	10	10	11
Syntactic fail	10	10	10	12

The overall results of these additional analyses are summarized in Table 5. Clearly, if we continue raising the frequency threshold, more derivations will fail, until eventually we will reach a point where no sentence will be derivable from units attested in the main corpus. However, the slope of the curve corresponding to the increase in fails is very gentle. This is because, as we have seen earlier, most of the component units used in the derivation are quite frequent. Perhaps surprisingly, raising the main corpus frequency threshold has very little effect on the final outcome: typically, the target sentences can still be derived from the component units, but may require additional operations. Note, too, that, in spite of the fact that most of the low frequency units are phrasal, raising the frequency threshold is equally likely to lead to a lexical fail or to a syntactic fail.

6. How adults build sentences

I have argued that ordinary language use, whether by adults or by children, involves "recycling" fragments of utterances by superimposition and juxtapositior, or what Tomasello (2000: 74) calls "cut-and-paste". It is important to emphasize that constructing novel utterances by combining chunks requires grammatical knowledge. To be able to produce well-formed utterances using this method, speakers must know which chunk of form corresponds to which chunk of meaning (i.e., have some understanding of constituency), and they must know which slots can be filled by which items (i.e., they must have some notion of grammatical categories such NP, VP, and VP, as well as more specific subcategories like "verb phrase with bare verb", "gerundive verb phrase", "verb phrase headed by past participle"). However, this kind of knowledge is rather different in kind from the abstract rules that, according to most linguists, are necessary to explain our linguistic productivity. Note, too, that the knowledge required for recycling can be construction specific. It is conceivable, for instance, that a particular speaker has passive constructions which apply to specific verbs (e.g. *NP1 is based on NP2, NP1 was known to NP2*, but not a verb-general passive construction (*NP1 was VERBed by NP2*). Categories, too, can be construction-specific, e.g. in *the SOUND of the N* construction, the first slot can be filled either by a noun designating sound (*sound, noise*) or by an onomatopoeic expression (*chuff@o,* *drip+drip@o*), while the second slot can be filled with any noun referring to a source of sound (e.g. *lorry, train, tap,* etc.)

Most earlier work on holistic processing regards it as a special mode, something that people use to complement the normal (i.e. analytic) mode (but see Sinclair 1991). My claim here is that recycling chunks is the *basic mode* of sentence production – what people normally do. Is this all there is to knowing a language? Almost certainly not. There is evidence that adults – particularly adults in societies with compulsory schooling and a long tradition of literacy – also have more abstract knowledge. As we have seen, some structures (for instance, relative clauses and some kinds of non-finite clauses) are difficult to accommodate in a lexically specific framework and may require more abstract representations. Note, however, that such structures are most typical of written language, not informal conversation. The basic "recycling" mode can be complemented by a more analytic mode; the latter, however, places much higher demands on the processing system and may be a product of formal schooling (Givón 1998; Iwasaki under review; Perfetti and McCutchen 1987; Wray and Grace 2007).

In fact, it may be possible to avoid the analytic mode altogether by applying analogy when no suitable lexically-specific schema is available. Dąbrowska (2008a) argues that English speakers have two lexically specific formulaic frames for questions with long distance dependencies: *WH do you think S-GAP?* and *WH did you say S-GAP?,* and that they produce "prototypical" questions with long-distance dependencies, i.e., questions that match one of these templates, by inserting a WH word into the first slot and a clause with a missing constituent into the second slot, and modifying the template (e.g. by changing the verb or the subject) when they want to produce questions such as *What do you believe he said?* or *What did he say she wanted?* If we accept this view, then the "analytic" mode may be analogous to what Moon (1998) calls "exploitations" of idioms such as *throw in the moist towelette* or *use* *an earthmover to crack a nut* (modelled on *throw in the towel* and *use a sledgehammer to crack a nut,* respectively).

What, then, would a lexically specific usage-based account of language production look like? I have argued that speakers possess vast inventories of fixed phrases and formulaic frames (and possibly some more general constructions as well) stored in a richly interconnected associative network. During production, they retrieve the constructions which partially match their communicative intentions and combine them by juxtaposing or superimposing them as appropriate. It is likely that speakers often retrieve more units than they actually need to construct their utterance, and so several different verbalizations of the message may be assembled in parallel. Normally, the verbalization which is completed first would be produced, although in some circumstances a speaker may reject it in favour of an alternative verbalization. I also assume that regular combinations are available as units if they are frequent enough. We do not know how frequent a complex expression must be in order to be available as a preconstructed unit, and it is likely that there is a trade-off between frequency and specificity: more frequent units may be more entrenched and hence easier to retrieve, but larger units show more overlap with the speakers' communicative intentions and result in simpler derivations. It is possible that different speakers resolve this conflict in different ways: some may prefer larger, more concrete units (and produce fluent though stereotypical utterances) while others may rely on smaller chunks. The degree to which a speaker relies on relatively large, concrete units also depends to a certain extent on the situation and/or topic: speakers may be very fluent in situations for which they have a large repertoire of phrases, and much less fluent in other contexts. This is nicely illustrated by Kuiper's (1996) study of abnormally fluent speakers such as auctioneers and fast sport commentators, as well as the common observation that bilinguals who are used to speaking different languages in different contexts may become surprisingly disfluent when required to use one of their languages in the "wrong" setting.

Note that the account proposed here differs from the traditional view which assumes that chunks are stored as unanalysed units. While this may be the case for some chunks, I propose that most chunks are analysed into their components. This is a precondition for our ability to recombine units to produce novel utterances, and it also ensures more efficient storage. Thus, the fact that *at the end of the day* is available as a unit does not prevent the language user from perceiving, and being able to access, the smaller chunks that make it up, including individual words (*at, the, end,* etc.), as well as intermediate-sized chunks such as *the end, the end of NP, the day,* and so on. This is a consequence of the way that memory works. When we commit something to memory – be it of a scene we have observed or an utterance we have heard – we extract key elements from the experience and encode them, with different parts of the experience stored separately in different parts of the brain. When we subsequently recall the experience, we access the components and reconstruct the memory rather than retrieving copies of the original experience (Fernyhough 2012; Fuster 1999).

Thus, I suggest that chunks of various sizes co-exist, not as distinct units, but as partially overlapping representations which may be thought of as a family of units whose members reinforce each other. A language user can access any of the subchunks independently, but can also use the smaller, more frequent chunks as retrieval cues for larger ones: in other words, a large and relatively infrequent unit such as *at the end of the day* may be accessed by activating the semantic units corresponding to END and DAY, which in turn activate their phonological representations and the larger symbolic unit of which they are both a part.

The “recycling” account of syntactic productivity proposed here has some similarities with lexically-based models of grammatical encoding such as Levelt's (1989) and especially the more recent approach developed by Ferreira (2000), in which individual lexical items are thought to include elementary trees which can be unified by substitution or adjoining. It differs from such models in two important ways. First, it is based on a cognitive linguistic framework which treats words as mini-constructions (cf. Dąbrowska 2009) rather than assuming autonomous syntactic representations. Secondly, drawing on the extensive literature on multiword chunks, it proposes the units that are combined are typically larger than simple lexical items.

In this respect, the recycling model resembles “Data-Oriented Parsing” (DOP), a computational approach developed by Rens Bod and his colleagues (Scha et al. 1999; Bod 2006, 2009). Like the approach outlined here, DOP is based on the principle that people store analyzed fragments of previously encountered sentences and recombine these fragments to produce novel sentences; however, since the stored fragments are parse trees, it assumes autonomous syntactic representations. In order to explain acquisition, DOP also makes another psychologically implausible assumption, namely, that the learner initially allows for all possible trees for a given sentence and decides on which one is correct on the basis of experience (Bod 2006, 2009). Since the number of possible trees grows exponentially with sentence length, it is difficult to see how this approach would avoid a computational explosion. The great advantage of DOP, on the other hand, is the fact that it is computationally implemented, thus making it relatively easy to test the effects of various assumptions.

Experiments with DOP have produced some interesting results which are very relevant to the proposal developed here. Scha et al. (1999) report that using larger, and in particular, overlapping fragments leads to better performance, although the returns are diminishing: thus while increasing fragment size from 1 to 2 and from 2 to 3 items results in a considerable increase in the model's accuracy, further increases result in only modest improvements, and there is no improvement for chunks of more than 6 items. (It is interesting to note in this connection that most of the chunks that were actually used in the traceback were 2-3 words long, and chunks of more than 6 words were extremely rare.) Furthermore, Scha et al. found that using lexically specific fragments also resulted in better performance – although again, only up to a point, with accuracy actually deteriorating when fragments were more than 8 words long – and that the use of low frequency fragments improved accuracy (primarily because low-frequency fragments tend to be large and lexically specific).

The "recycling" model outlined here provides a simple and psychologically plausible account of language production in informal conversation in English, a language with a relatively fixed word order. Could it also be applied to languages with a richer morphology and more variable word order? It is safe to assume that speakers of such languages store fixed phrases as well as some frames with slots: very few, if any, languages have no restrictions on word order at all, so some frames will be needed to explain constructions with fixed word order. However, to explain constructions which allow variable word order, we need a new type of unit, which I will call a **packet**: a multi-word unit which does not specify the order of the components. Thus, while an English transitive verb is a frame with two slots (*NP1 > verb > NP2*), a transitive verb in a relatively free word order language such as Polish would be a packet consisting of a verb and two nominals (*verb, NP1, NP2*): note that I am using they symbol '>' to indicate ordering and the comma to indicate lack of ordering. The units contained in a packet would be juxtaposed when the sentence is actually produced. Statistical preferences for certain word orders would be explained partly by more general word order constructions and partly by processing constraints: for instance, the fact that speakers tend to produce the most accessible elements first in order to free working memory (Ferreira and Engelhardt 2006) can explain the preference for ordering given information before new.

As indicated earlier, phrase recycling can easily accommodate grammatical relationships such as government and agreement if we assume that words are partially schematic constructions: for instance, superimposing *(NP) eat NP* and *V them* (rather than just the simple pronoun) ensures that the appropriate form of the pronoun appears after the verb; superimposing a schematic nominal construction (*two N-PL*) and a plural noun (*>1 mice*) ensures agreement between the nominal and the noun, and so on. In languages where the grammatical relationship is signaled on the dependent word, the larger construction will specifify details of dependent marking. For instance, in Polish, the object of a negated verb is marked by the genitive case:

(4) *Nie ma soku.*
not have:3SG juice:GEN:SG
'There isn't any juice.'

Assuming that the packet [*nie ma, N-u*] 'there isn't any N' and the lexically specific unit *sok-u* 'juice:GEN' are available to the speaker as chunks, sentence (4) would be assembled by juxtaposing the component units of the packet and superimposing *sok-u* and (*nie ma) N-u.* Clearly, a detailed analysis will be necessary to determine whether such a method would actually work; hopefully, however, this brief discussion demonstrates how it could work in principle.

7. Beyond productivity

The account outlined above raises some interesting questions. If different speakers have different lexically specific chunks, how can they understand each other? And in what sense can they be said to speak the same language? Last but not least, languages are not just collections of chunks: they contain numerous regularities. Where do these regularities come from?

As we saw earlier, the same utterances can be assembled using different chunks; thus the fact different speakers store different lexically specific units is not necessarily a problem for this account. (Morevoer, even the same speaker may assemble the same utterance using different chunks on different occasions, depending on, for example, which units have been primed by prior discourse. This flexibility helps to explain the speed of language processing: we save time by opportunistically using whichever chunks are most accessible at the time of the speech event.) Note, too, that speakers don't need to share exactly the same grammar to be able communicate. Human beings are very good at using whatever information is available to infer other people's communicative intentions: indeed this is what makes language acquisition possible in the first place (Tomasello 2008). Furthermore, as Ferreira et al. (2002) note, in most casual conversations, a very superficial analysis of the linguistic form is "good enough" for successful communication (see also Dąbrowska 2004, chapter 2).

Secondly, in what sense can speakers with (possibly very) different inventories of lexically specific chunks be said to "speak the same language"? The obvious answer is that chunks *are* shared by speakers, although they need not be shared by all speakers. An analogy with vocabulary may be helpful here. Vocabulary is shared, in the sense that (presumably) every English speaker knows words like *good, boy,* or *sleep*; many speakers also know words like *bawdy, trudge,* or *ubiquitous*; however, very few have productive knowledge of words like *corpuscle* or *dunnart* – and yet all of these words are considered part of the language. Complex units are shared in the same way. Highly frequent units such as those in (5) are shared by (nearly) everyone, while less frequent units are shared by only a few speakers. Languages belong to communities, not to individual speakers: an individual speaker "owns" only a part of his or her language.

(5) *I don't think so.* (359)

I don't believe a word of it. (78)

I haven't seen you for ages. (101)

Who do you think you are? (531)

What's a nice girl like you doing in a place like this? (157)[1]

Language is also shared in a different sense. Every time a speaker produces an utterance and the listener understands his or her communicative intentions, that particular form-meaning pairing is shared by them, at least momentarily. In other words, every successful act of communication is an act of sharing linguistic expressions, and any linguistic device that contributed to communicative success is reinforced. As Millikan (2008) argues, a linguistic convention will survive if it succeeds in coordinating speaker-hearer communicative intentions *some critical proportion of the time:* it doesn't have to be all of the time, or even most of the time.

Finally, why are languages as regular as they are? Before we address this question, it is perhaps worth noting that modern linguistics has tended to overemphasize the systematic nature of language, and overlook the idiosyncrasies. It is undeniable that languages exhibit many regularities – but they are also shot through with irregularities: as Sapir famously observed, "all grammars leak" (1921: 39). Furthermore, standard varieties of modern

[1] The numbers in parentheses indicate WebCorp frequencies. The search was carried out on 2 March 2013).

“exoteric” languages, which are the subject of the most intensive linguistic enquiry, are products of a long history of schooling and language planning, and thus tend to be more regular than languages in their “natural” state (see Wray and Grace 2007).

The traditional explanation for systematicity is that all languages are highly constrained by an innate Universal Grammar, or alternatively, that speakers prefer more general rules. As I argue elsewhere (Dąbrowska 2010, under review), neither of these explanations is very satisfactory. An alternative account is provided by functionalist studies which demonstrate that many patterns in language arise as a result of processing demands (Hawkins 2004; Jaeger and Tily 2011; MacDonald 2013), communicative pressures having to do with economy, i.e., concise representation of frequent concepts (Bybee 2010; Haspelmath 2008, 2014; Jaeger and Tily 2011) on the one hand, and avoiding ambiguity (Gibson et al. 2013) on the other, as well as a variety of other discourse and cognitive pressures (Ariel 2009, Du Bois 2003; Givón 2009). Such pressures can be quite weak: all they have to do is increase the frequency of one variant slightly. Once this has occurred, frequency itself will tend to amplify the pattern, since more entrenched patterns are easier to access and hence preferred by speakers, which results in even greater frequency, and so on. Finally, regular patterns can arise as a result of intergenerational transmission. This is nicely demonstrated by a series of experiments on “iterated learning” in which speakers are exposed to a set of form-meaning pairings and subsequently produce them; their output serves as input for the next “generation” of participants, who in turn provide input for the next group, and so on (see Cornish et al. 2009; Scott-Phillips and Kirby 2010). The form-meaning pairings are initially random, but gradually evolve into a system. Crucially, this happens not because learners necessarily have a preference for more regular mappings, but simply as a result of the fact that each learner is exposed to only a subset of the possible form-meaning pairings:

> Over time, iterated learning ensures languages evolve to (a) become easier to learn and (b) become more structured….. This is because there are only two ways to survive the transmission bottleneck: be heard (and remembered) by the next generation or be easily inferable from what is heard. This latter solution can only occur when there is some kind of regularity to be exploited in the system. (Cornish et al. 2009: 188-189)

Traditionally, linguists have assumed that speakers of the same language variety share the same grammar, and that speakers' grammars are the same as linguists' grammars. There is growing evidence that this is not the case: unlike linguists, speakers appear to have a preference for low-level local schemas rather than general rules which apply "across the board" (Dąbrowska 2010), and there are substantial differences between individual speakers' grammars (Dąbrowska 2012). However, speakers attempt to approximate each other's behaviour, and thus community grammars appear to be more regular than individual grammars (Dąbrowska 2008b, 2013; Hurford 2000). It follows that there are regularities in language which are not explicitly represented in speakers' minds (or at least not all speakers' minds). To understand how such regularities arise, we must learn to think of language as complexity theory sees it: as a complex, multi-agent, dynamic and adaptive system where regularities emerge from interactions between speakers (Larsen-Freeman and Cameron 2008) and are best viewed as central tendencies in a massive collection of previously experienced utterances (Ellis and Larsen-Freeman 2006).

References

Altenberg, Bengt. 1998. On the phraseology of spoken English: The evidence of recurrent word-combinations. In A. P. Cowie (ed.). *Phraseology: Theory, analysis and applications,* 101-122. Oxford: Clarendon Press.

Ariel, Mira. 2009. Discourse, grammar, discourse. *Discourse Studies* 11. 5-36.

Arnon, Inbal & Neal Snider. 2010. More than words: Frequency effects for multi-word phrases. *Journal of Memory and Language* 62. 67-82.

Barlow, Michael & Suzanne Kemmer. 2000. *Usage-based models of language*. Cambridge: Cambridge University Press.

Barsalou, Lawrence W. 1992. *Cognitive psychology: An overview for cognitive scientists*. Hillsdale, NJ: Lawrence Erlbaum.

Berman, Ruth A. 2007. Developing linguistic knowledge and language use across adolescence. In Erika Hoff & Marilyn Shatz (eds.). *Blackwell handbook of language development*. Oxford: Blackwell Publishing.

Berman, Ruth A. (ed.). 2004. *Language development across childhood and adolescence*. Amsterdam: Benjamins.

Bod, Rens. 2006. Exemplar-based syntax: How to get productivity from examples. *The Linguistic Review* 23. 291-320.

Bod, Rens. 2009. From exemplar to grammar: A probabilistic analogy-based model of language learning. *Cognitive Science* 33. 752-793.

Bowerman, Melissa. 1996. The origins of children's spatial semantic categories: cognitive versus linguistic determinants. In John J. Gumperz & Stephen C. Levinson (eds.). *Rethinking linguistic relativity*, 145-176. Cambridge: Cambridge University Press.

British National Corpus, The. Version 2 (BNC World). 2001. BNC Consortium.

Bull, Matthew & Matthew Aylett. 1998. An analysis of the timing of turn-taking in a corpus of goal-orientated dialogue *Proceedings of ICSLP-98, Sidney, Australia*. (Vol. 4), 1175-1178.

Bybee, Joan. 2010. *Language, usage and cognition*. Cambridge: Cambridge University Press.

Cameron-Faulkner, Thea, Elena Lieven, & Michael Tomasello. 2003. A construction based analysis of child-directed speech. *Cognitive Science* 27. 843-873.

Conklin, Kathy & Norbert Schmitt. 2008. Formulaic sequences: Are they processed more quickly than nonformulaic language by native and nonnative speakers? *Applied Linguistics* 29. 72-89.

Cornish, Hannah, Monica Tamariz & Simon Kirby. 2009. Complex adaptive systems and the origins of adaptive structure: what experiments can tell us. *Language Learning* 59: Suppl. 1. 187-205.

Cowie, A. P. 1998. *Phraseology*. Oxford: Clarendon Press.

Dąbrowska, Ewa. 2000. From formula to schema: The acquisition of English questions. *Cognitive Linguistics* 11. 83-102.

Dąbrowska, Ewa. 2004. *Language, mind and brain. Some psychological and neurological constraints on theories of grammar*. Edinburgh: Edinburgh University Press.

Dąbrowska, Ewa. 2008a. Questions with 'unbounded' dependencies: A usage-based perspective. *Cognitive Linguistics* 19. 391-425.

Dąbrowska, Ewa. 2008b. The later development of an early-emerging system: The curious case of the Polish genitive. *Linguistics* 46. 629-650.

Dąbrowska, Ewa. 2009. Words as constructions. In Vyvyan Evans & Stéphanie Pourcel (eds.). *New directions in cognitive linguistics*, 201-223. Amsterdam: John Benjamins.

Dąbrowska, Ewa. 2010. The mean lean grammar machine meets the human mind: Empirical investigations of the mental status of rules. In Hans-Jörg Schmid and Susanne Handl (eds.). *Cognitive foundations of linguistic usage patterns. Empirical approaches*, 151-170. Berlin: Mouton de Gruyter.

Dąbrowska, Ewa. 2012. Different speakers, different grammars: Individual differences in native language attainment. *Linguistic Approaches to Bilingualism* 2. 219-253.

Dąbrowska, Ewa. 2013. Functional constraints, usage, and mental grammars: A study of speakers' intuitions about questions with long-distance dependencies. *Cognitive Linguistics* 24. 633 - 665.

Dąbrowska, Ewa. Under review. What exactly is Universal Grammar, and has anyone seen it?

Dąbrowska, Ewa & Elena Lieven. 2005. Towards a lexically specific grammar of children's question constructions. *Cognitive Linguistics* 16. 437–474.

Dąbrowska, Ewa, Caroline Rowland & Anna Theakston. 2009. The acquisition of questions with long-distance dependencies. *Cognitive Linguistics* 20. 571-596.

Diessel, Holger. 2004. *The Acquisition of Complex Sentences*. Cambridge: Cambridge University Press.

Du Bois, John W., Wallace L. Chafe, Charles Meyer, Sandra A. Thompson, Robert Englebretson & Nii Martey. 2000-2005. Santa Barbara Corpus of Spoken American English, Parts 1-4. Retrieved 20 March 2014, from http://www.linguistics.ucsb.edu/research/santa-barbara-corpus.

Du Bois, John W. 2003. Discourse and grammar. In Michael Tomasello (ed.). *The new psychology of language: Cognitive and functional approaches to language structure* (Vol. 2), 47-87. London: Erlbaum.

Ellis, Nick C. 1996. Phonological memory, chunking, and points of order. *Studies in Second Language Acquisition* 18. 91-126.

Ellis, Nick C. & Diane Larsen-Freeman. 2006. Language emergence: Implications for Applied Linguistics -- Introduction to the special issue. *Applied Linguistics* 27. 558-589.

Erman, Britt & Beatrice Warren. 2000. The idiom principle and the open choice principle. *Text* 20. 29-62.

Fernyhough, Charles. 2012. *Pieces of light: The new science of memory*. London: Profile Books.

Ferreira, Fernanda. 2000. Syntax in language production: An approach using tree-adjoining grammars. In LindaWheeldon (ed.). *Aspects of language production*, 291-330. Philadelphia: Psychology Press.

Ferreira, Fernanda & Paul E. Engelhardt. 2006. Syntax and production. In Matthew Traxler & Morton Ann Gernsbacher (eds.). *Handbook of Psycholinguistics*. New York: Academic Press.

Ferreira, Fernanda, Karl G. D. Bailey & Vittoria Ferraro. 2002. Good-enough representations in language comprehension. *Current Directions in Psychological Science* 11(1). 11-15.

Fuster, Joaquin M. 1999. *Memory in the cerebral cortex. An empirical approach to neural networks in the human and nonhuman primate*. Cambridge, MA: MIT Press.

Gibson, Edward, Piantadosi, Steven T., Brink, Kimberly, Bergen, Leon, Lim, Eunice & Saxe, Rebecca 2013. A noisy-channel account of cross-linguistic word order variation. *Psychological Science* 24. 1079-1088.

Givón, Talmy 1998. The grammar of literacy. *Syntaxis* 1. 1-98.

Givón, Talmy 2009. *The genesis of syntactic complexity: Diachrony, pntogeny, neuro-cognition, evolution*. Amsterdam: John Benjamins.

Goldberg, Adele E. 1995. *Constructions. A Construction Grammar approach to argument structure*. Chicago: The University of Chicago Press.

Goldberg, Adele E. 2006. *Constructions at work. The nature of generalization in Language*. Oxford: Oxford University Press.

Gurevich, Olya, Matthew A. Johnson & Adele E. Goldberg. 2010. Incidental verbatim memory for language. *Language and Cognition* 2. 45-78.

Haspelmath, Martin. 2008. Frequency vs. iconicity in explaining grammatical asymmetries. *Cognitive Linguistics* 19. 1-33.

Haspelmath, Martin. 2014. On system pressure competing with economic motivation. In Brian MacWhinney, Andrej L. Malchukov & Edith A. Moravcsik (eds.). *Competing motivations*. Oxford: Oxford University Press.

Hawkins, John A. 2004. *Efficiency and complexity in grammars*. Oxford: Oxford University Press.

Hirsh-Pasek, Kathy & Roberta Michnick Golinkoff. 1996. *The origins of grammar: Evidence from early language comprehension*. Cambridge, MA: MIT Press.

Hoff, Erika. 2009. Language development at an early age: Learning mechanisms and outcomes from birth to five years *Encyclopedia of early childhood Development*. Available from http://www.child-encyclopedia.com/en-ca/language-development-literacy/according-to-experts/hoff.html.

Hoffer, Bates L. 1994. The last stages of the acquisition of syntax: The interaction of reading and syntactic development. In Valerie Becker Makkai (ed.). *The twentieth LACUS Forum 1993*, 654-663. Chapel Hill, NC: Linguistic Association of Canada and the United States.

Hurford, James R. 2000. Social transmission favours linguistic generalisation. In Chris Knight, Michael Studdert-Kennedy & James R. Hurford (eds.). *The evolutionary* *emergence of language: Social function and the origins of linguistic form*, 324–352. Cambridge: Cambridge University Press.

Iwasaki, Shoichi. Under review. A multiple-grammar model of speakers' linguistic knowledge.

Jaeger, T. Florian & Tily, Harry. 2011. On language ‘utility’: Processing complexity and communicative efficiency. *Wiley Interdisciplinary Reviews: Cognitive Science* 2. 323–335.

Johansson, Victoria. 2008. Lexical diversity and lexical density in speech and writing: a developmental perspective. *Working Papers, Department of Linguistics and Phonetics, Lund University* 53. 61-79.

Kuiper, Koenraad. 1996. *Smooth talkers. The linguistic performance of auctioneers and sportscasters*. Mahwah, NJ: Lawrence Erlbaum.

Langacker, Ronald W. 1987. *Foundations of Cognitive Grammar. Volume 1: Theoretical prerequisites*. Stanford, CA: Stanford University Press.

Langacker, Ronald W. 2000. A dynamic usage-based model. In Michael Barlow & Suzanne Kemmer (eds.). *Usage-based models of language*, 1-63. Stanford, CA: CSLI Publications.

Larsen-Freeman, Diane & Lynne Cameron. 2008. *Complex systems and applied linguistics*. Oxford: Oxford University Press.

Levelt, Willem J. M. 1989. *Speaking: From intention to articulation*. Cambridge, MA. MIT Press.

Lieven, Elena V., Heike Behrens, Jennifer Speares & Micheal Tomasello. 2003. Early syntactic creativity: A usage-based approach. *Journal of Child Language* 30. 333-370.

Lieven, Elena V., Julian M. Pine & Gillian Baldwin. 1997. Lexically-based learning and early grammatical development. *Journal of Child Language* 24. 187-219.

Lieven, Elena V., Dorothé Salomo & Michael Tomasello. 2009. Two-year-old children's production of multiword utterances: A usage-based analysis. *Cognitive Linguistics* 20. 481-508.

Logan, Gordon D. 1988. Toward an instance theory of automatization. *Psychological Review* 95. 492-527.

MacDonald, Maryellen C. 2013. How language production shapes language form and comprehension. *Frontiers in Psychology, 4*, 226.

MacWhinney, Brian. 2014. *The CHILDES project. Tools for analyzying talk – (Electronic edition) Part 2: The Clan Programs.* Retrieved from http://childes.psy.cmu.edu/manuals/clan.pdf

Millikan, Ruth G. 2008. A difference of some consequence between conventions and rules. *Topoi* 27. 87-99.

Moon, Rosamund 1998. *Fixed expressions and idioms in English: A corpus-based approach*. Oxford: OUP.

Newport, Elissa L.,Henry Gleitman & Lila R. Gleitman. 1977. Mother, I'd rather do it myself: some effects and non-effects of maternal speech style. In Catheirne E. Snow & Charles A. Ferguson (eds.). *Talking to children. Language input and acquisition*, 109-149. Cambridge: Cambridge University Press.

Nippold, Marilyn A. 1998. *Later language development: The school-age and adolescent years*. Austin, Texas: Pro-ed.

Nippold, Marilyn A., Linda J. Hesketh, Jill K. Duthie & Tracy C. Mansfield. 2005. Conversational v. expository discourse: A study of syntactic development in children, adolescents and adults. *Journal of Speech, Language and Hearing Research 48*. 1048-1064.

Pawley, Andrew & Frances H. Syder. 1983. Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In Jack C. Richards & Richard W. Schmidt (eds.). *Language and Communication*, 191-226. London: Longman.

Perfetti, Charles A. & Deborah McCutchen. 1987. Schooled language competence: linguistic abilities in reading and writing. In Sheldon Rosenberg (ed.). *Advances in applied psycholinguistics. Volume 2: Reading, writing, and language learning*, 105-141. Cambridge: Cambridge University Press.

Pinker, Steven 1995. Language acquisition. In Lila R. Gleitman & Mark Liberman (eds.), *An invitation to Cognitive Science*. *Volume 1: Language* (2nd edition), 135-182. Cambridge, MA: MIT Press.

Pollack, Irwin & Pickett, J. M. 1964. Intelligibility of excerpts from fluent speech: auditory vs. structural context. *Journal of Verbal Learning and Verbal Behavior* 3. 79-84.

Richards, Brian J. & David D. Malvern. 1997. *Quantifying Lexical Diversity in the Study of Language Development*. Reading: Faculty of Education and Community Studies, University of Reading.

Sapir, Edward. 1921. *Language: An introduction to the study of speech*. New York: Harcourt, Brace and Company.

Scha, Renko, Rens Bod & Khalil Sima'an. 1999. A memory-based model of syntactic analysis: Data-oriented parsing. *Journal of Experimental and Theoretical Artificial Intelligence* (Special Issue on Memory-Based Language Processing) 11. 409-440.

Schneider, Wolfganng & David F. Bjorklund. 1998. Memory. In Deanna Kuhn & Robert S. Siegler (eds.). *The handbook of child psychology. Volume 2: Cognition, perception and language*, 467-522. New York: John Wiley and Sons.

Scott-Phillips, Thom C. & Simon Kirby. 2010. Language evolution in the laboratory. *Trends in Cognitive Sciences* 14. 411-417.

Shepard, Roger N. 1967. Recognition memory for words, sentences and pictures. *Journal of Verbal Learning and Verbal Behavior* 6. 156-163.

Sinclair, John. 1991. *Corpus, concordance, collocation*. Oxford: Oxford University Press.

Sinha, Chris, Lis A. Thorseng, Mariko Hayashi & Kim Plunkett. 1994. Comparative spatial semantics and language acquisition: Evidence from Danish, English, and Japanese. *Journal of Semantics* 11. 253-287.

Siyanova-Chanturia, Anna & Ron Martinez (2014). The idiom principle revisited. *Applied Linguistics*.

Siyanova-Chanturia, Anna, Kathy Conklin, Walter J. B. van Heuven. 2011. Seeing a phrase "time and again" matters: The role of phrasal frequency in the processing of multi-word sequences. *Experimental Psychology: Learning, Memory, and Cognition* 37. 776-784.

Snow, Catherine. 1986. Conversations with children. In Paul Fletcher & Michael Garman (eds.). *Language Acquisition* (2nd edition), 69-89. Cambridge: Cambridge University Press.

Standing, Lionell. 1973. Learning 10,000 pictures. *Quarterly Journal of Experimental Psychology* 25. 207-222.

Stoll, Sabine, Kirsten Abbot-Smith & Elena V. Lieven. 2009. Lexically restricted utterances in Russian, German, and English Child-Directed Speech. *Cognitive Science* 33. 75–103.

Tauroza, Steve & Desmond Allison. 1990. Speech rates in British English. *Applied Linguistics* 11. 90-105.

Tomasello, Michael (2000). Do young children have adult syntactic competence? *Cognition* 74. 209-253.

Tomasello, Michael. 2003. *Constructing a language: A usage-based theory of child language acquisition*. Cambridge, MA: Harvard University Press.

Tomasello, Michael. 2008. *Origins of human communication*. Cambridge, MA: MIT Press.

Tremblay, Antoine & Benjamin V. Tucker. 2011. The effects of N-gram probabilistic measures on the recognition and production of four-word sequences. *The Mental Lexicon* 6. 302-324.

Tremblay, Antoine, Bruce Derwing, Gary Libben & Chris Westbury 2011. Processing advantages of lexical bundles: Evidence from self-paced reading and sentence recall tasks. *Language Learning* 61. 569-613.

Weinert, Regina. 2010. Formulaicity and usage-based language: linguistic, psycholinguistic and acquisitional manifestations. In David Wood (ed.). *Perspectives on formulaic language. Acquisition and communication*, 1-20. London/New York: Continuum.

Wray, Alison 2002. *Formulaic language and the lexicon*. Cambridge: Cambridge University Press.

Wray, Alison & George W. Grace 2007. The consequences of talking to strangers: Evolutionary corollaries of socio-cultural influences on linguistic form. *Lingua* 117. 543-578.

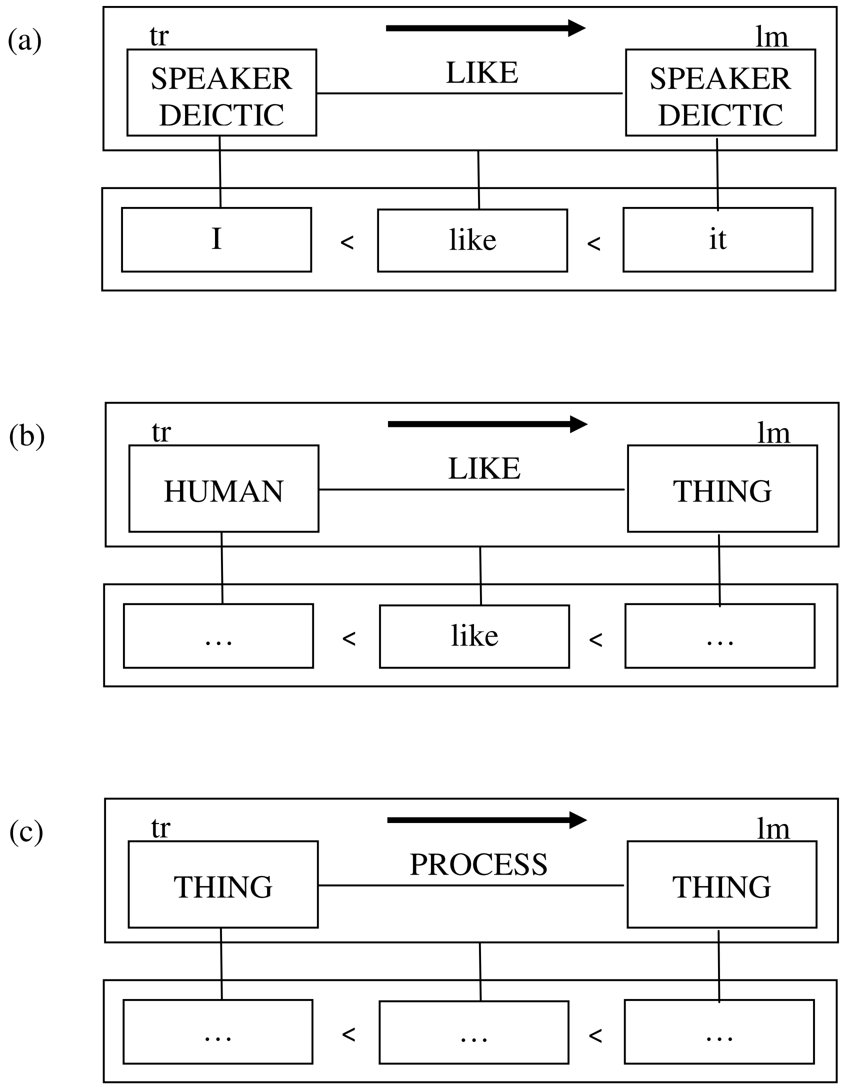

Figure 1: Types of complex units
(a) a formulaic (fixed) phrase
(b) a low level schema (“frame with slots”)
(c) a general schema (“abstract construction”)

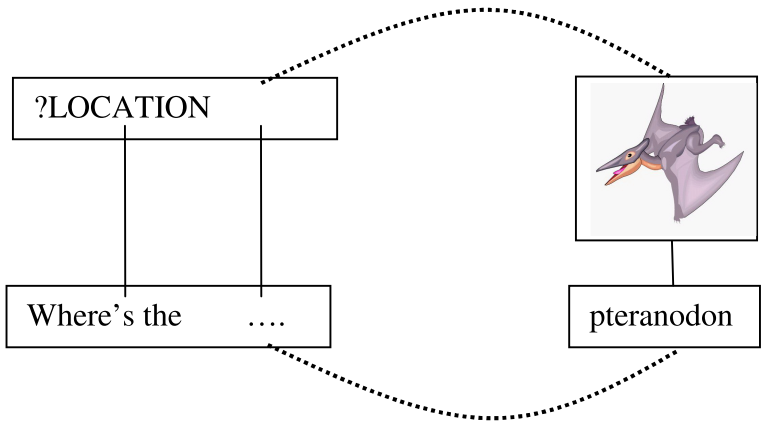

Figure 2: The superimposition of *where's the N* and *pteranodon*

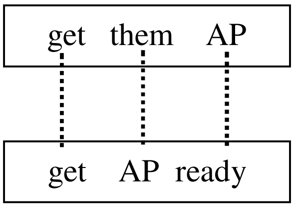

Figure 3: Superimposition of *get them AP* and *get AP ready*

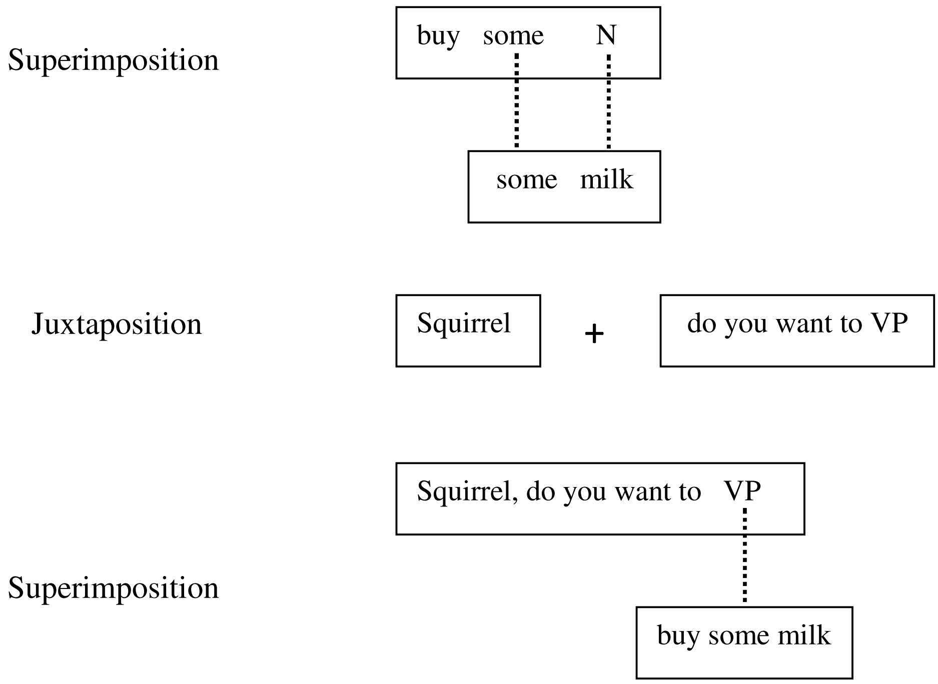

Figure 4: Derivation of Squirrel, do you want to buy some milk?

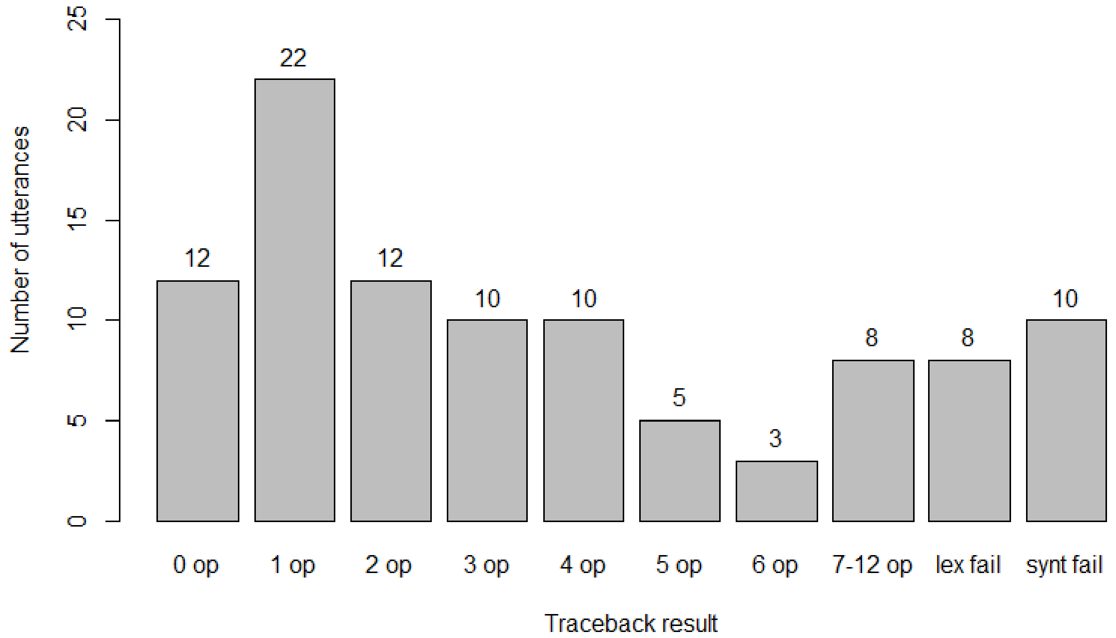

Figure 5: Number of operations needed to derive M’s utterances

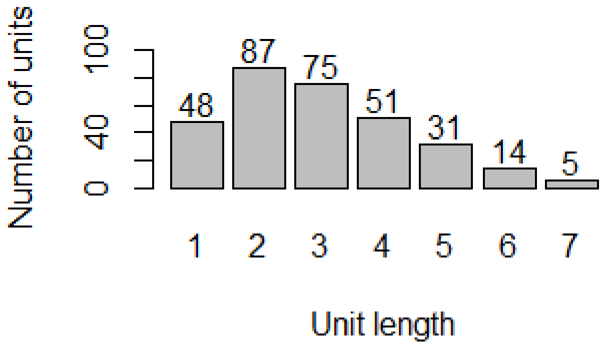

Figure 6: Frequency distribution of component units of different length

Notes

* Address for correspondence: E. Dąbrowska, Department of Humanities, Faculty of Arts, Design and Social Sciences, Northumbria University, Newcastle upon Tyne, NE1 8ST, United Kingdom. I would like to thank Kevin McManus for his help in analysing the data, and Vitor Zimmerer, Elena Lieven, Holger Diessel, and three anonymous *Cognitive Linguistics* referees for their comments on an earlier draft of this manuscript.

[i] Pollack and Pickett (1964) spliced single words out of a recorded conversation and played them to new participants. Only about 50% of the words were intelligible when presented in isolation.

[ii] The expression *I like it* occurs 389 times in the BNC, and thus is comparable to polysyllabic words such as *rectangle* (304 tokens in BNC), *erratic* (392 tokens) or *hideous* (423 tokens).

[iii] Consider, for example, the following statements:

> It is safe to say that except for constructions that are rare, predominantly used in written language, or mentally taxing even to an adult (like *The horse that the elephant tickled kissed the pig*), all parts of all languages are acquired before the child turns four. (Pinker 1995)

> By age 5, children essentially master the sound system and grammar of their language and acquire a vocabulary of thousands of words…. The development of complex (i.e., multi-clause) sentences usually begins some time before the child's second birthday and is largely complete by age 4. (Hoff 2009)

It is well known, however, that children acquire most of their grammar by the time they are three years old. (Hirsch Pasek & Golinkoff 1996: 2)

[iv] Note that Moon investigated only non-compositional expressions.

[v] One may ask whether this is indeed a single utterance or two, or perhaps even three. Spoken language researchers use a variety of criteria (e.g. pauses, intonation, syntactic integrity) to identify utterances – but these are often in conflict. Since CHAT (the transcription conventions employed in the CHILDES database) requires that each utterance be transcribed on a single line, for the purposes of this analysis, it was assumed that each line constituted a single utterance: in other words, I simply followed the intuitions of the original transcriber. Note that dividing such doubtful cases into shorter utterances would result in simpler derivations.

[vi] As explained earlier, in normal adult conversation about 210 words are spoken per minute. If the mother is 25 years old and was exposed to language for 6 hours a day all her life, her total linguistic experience comprises 689,850,000 words (210 × 60 × 6 × 365 × 25).